Mesoscale atmospheric reanalysis IPRA using 3DVAR over the Iberian Peninsula

Jon Sáenz
Depto. Física Aplicada II and PIE
UPV/EHU
http://www.ehu.es/eolo
Barcelona, 2014

Objectives

- To create a high-resolution WRF-based integration over the Iberian Peninsula using 3DVAR data assimilation.
  - Study processes
  - Fitting of statistical downscaling models
  - Generate an “as homogeneous as posible” reference dataset.
  - IPRA06: 1960-2002
- To evaluate the impact of data assimilation.
  - IP-NODA: WRF integration without data assimilation
  - IPRA-12: Assimilation every 12h (00Z, 12Z)
  - IPRA-06: Assimilation every 06h (00Z, 06Z, 12Z, 18Z)

Data

- Boundary and initial conditions: ERA40 (2D and 3D), GRIB, 1.125ºx1.125º
- Observations:
  - *Analysis feedback* (AF) from ERA40 (ECMWF *QC code: "used" in the AF files*)
  - Types used: TEMP, SYNOP, ACARS, SHIP, …
  - Types **not** used: TOVS, ATOVS, ...
- Verification: Daily data PRE, TEMP:
  - E-OBS 0.25º
  - SPAIN02
  - AEMET atlas, climatology 1971-2000
- Period:
  - NODA: 1990-01-01 to 1999-12-31
  - IPRA12: 1990-01-01 to 1999-12-31
  - IPRA06: 1960-01-01 to 2002-08-31

Method - NODA

- Integration (WRF 3.1.1) during 24h segments.
- Initialization first segment.
- Restart from model state every 24 hours.
- Parameterizations:
  - Microphysics WRF-5class
  - Radiation CAM (SW and LW)
  - 5-layer diffusive soil
  - PBL Mellor-Yamada-Janjic
  - Cumulus convection Betts-Miller-Janjic.
  - Surface: eta-similarity
- 35 vertical model levels, 15x15 km horizontal resolution

Method - WDA12h

- Integration (WRF 3.1.1) during 24h
- Stop at 12 hours
- Restart after 3DVAR 00Z and 12Z $t_{OBS} \in [t_{ANL} - 1\text{h}, t_{ANL} + 1\text{h}]$
- Background Error Covariances:
  - Seasonally varying (DecJanFeb->Jan, JFM->F ...).
  - “CV5” (WRFDA) method. Created ad-hoc with an initial 13 months integration (1990-01 to 1991-02).
- Parameterizations: Same as NODA
- Horizontal resolution and vertical levels same as in NODA

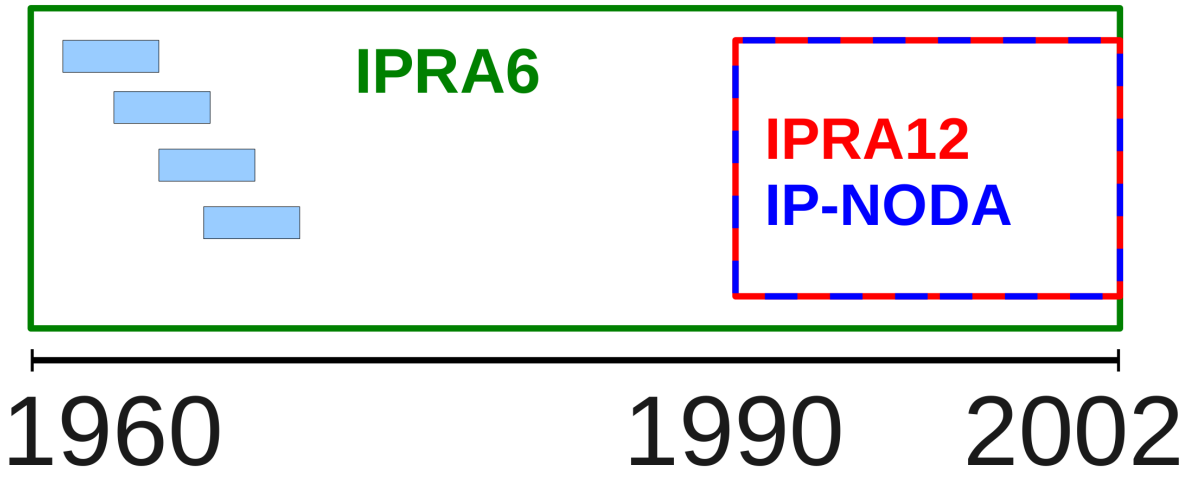
Method - WDA06h

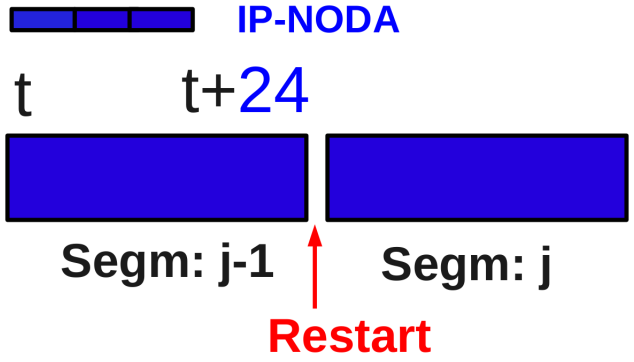
- Integration (WRF 3.1.1) during 12h
- Stop at 6 hours
- Restart after 3DVAR 00Z, 06Z, 12Z and 18Z

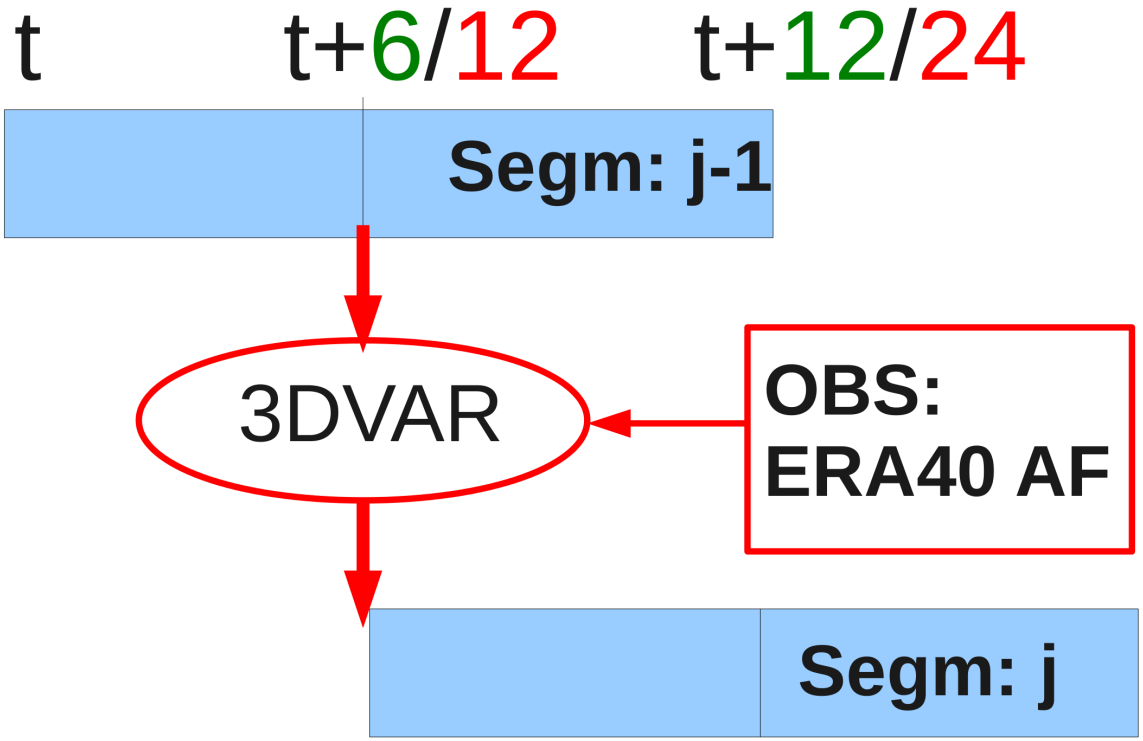
$$t_{OBS} \in [t_{ANL} - 1\text{h}, t_{ANL} + 1\text{h}]$$

- Background Error Covariances:
  - Seasonally varying (DecJanFeb->Jan, JFM->F ...).
  - “CV5” (WRFDA) method. Created ad-hoc with an additional 13 months integration (1990-01 to 1991-02).
- Parameterizations: Same as NODA, WDA12h
- Horizontal resolution and vertical levels same as in NODA, WDA12h

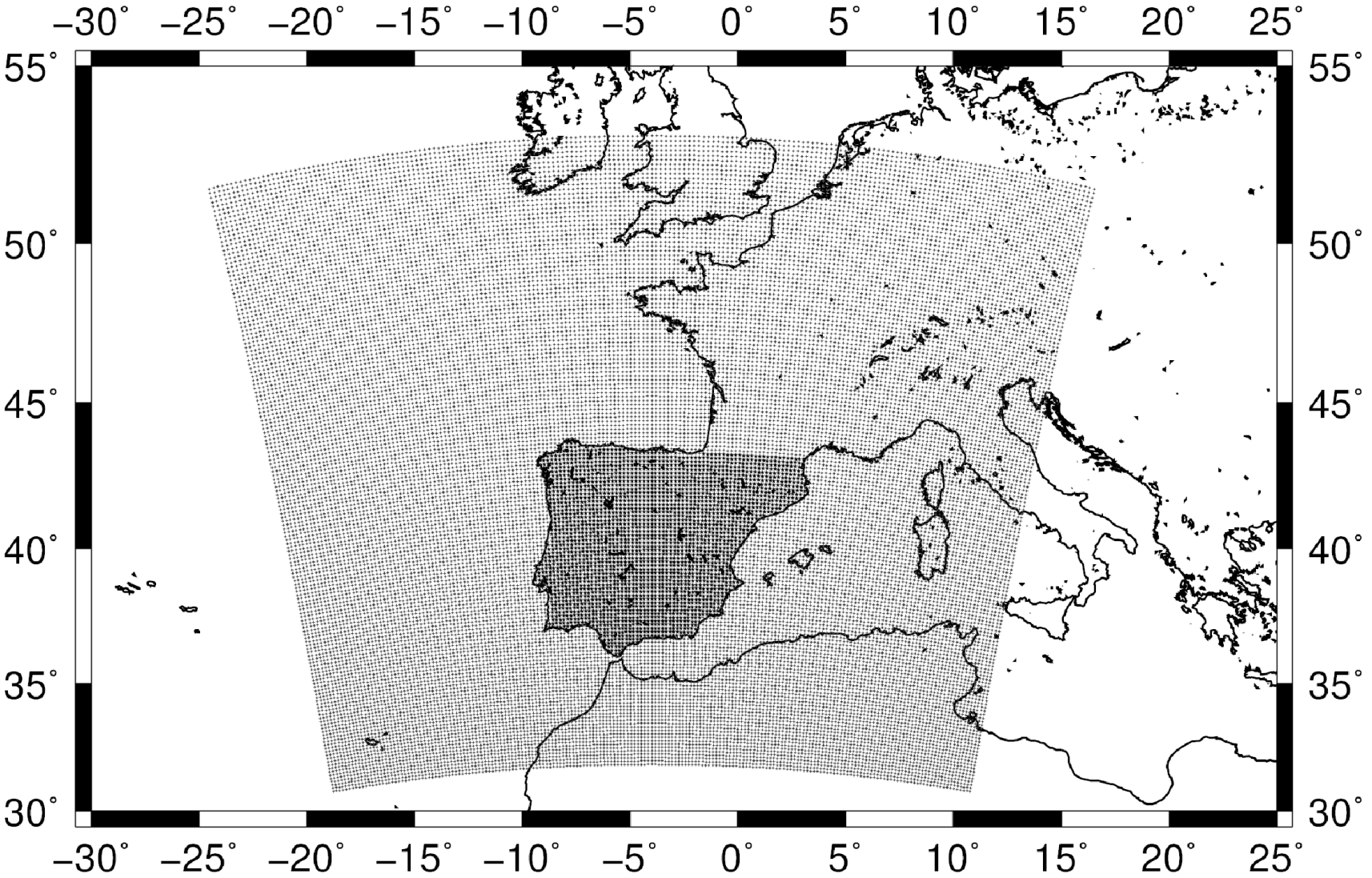
Schematic


IPRA6
IPRA12
IP-NODA
1960
1990
2002

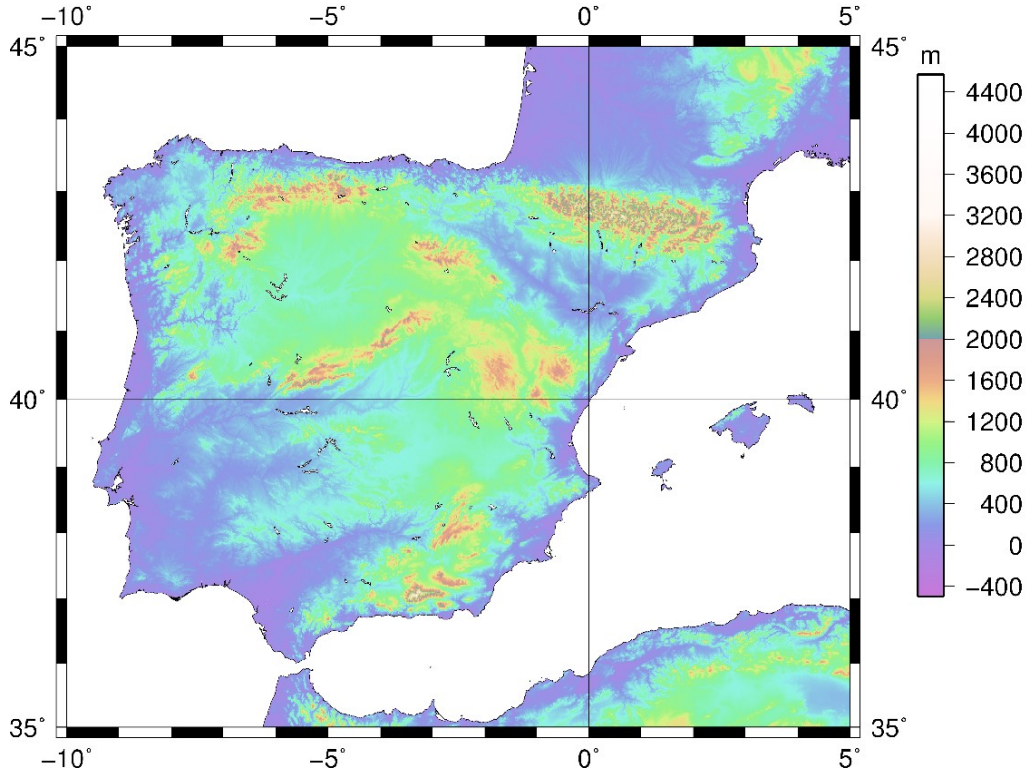

IP-NODA
t
t+24
Segm: j-1
Segm: j
Restart


t
t+6/12
t+12/24
Segm: j-1
3DVAR
OBS:
ERA40 AF
Segm: j

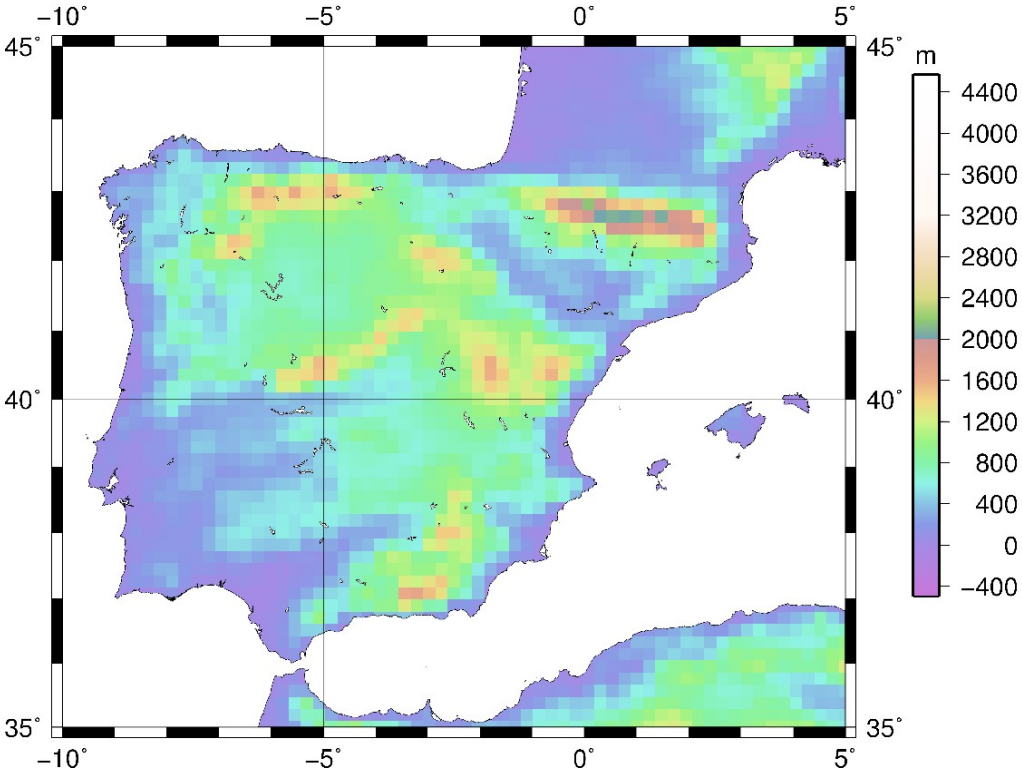
Domain and verification mask



Topography of the mask

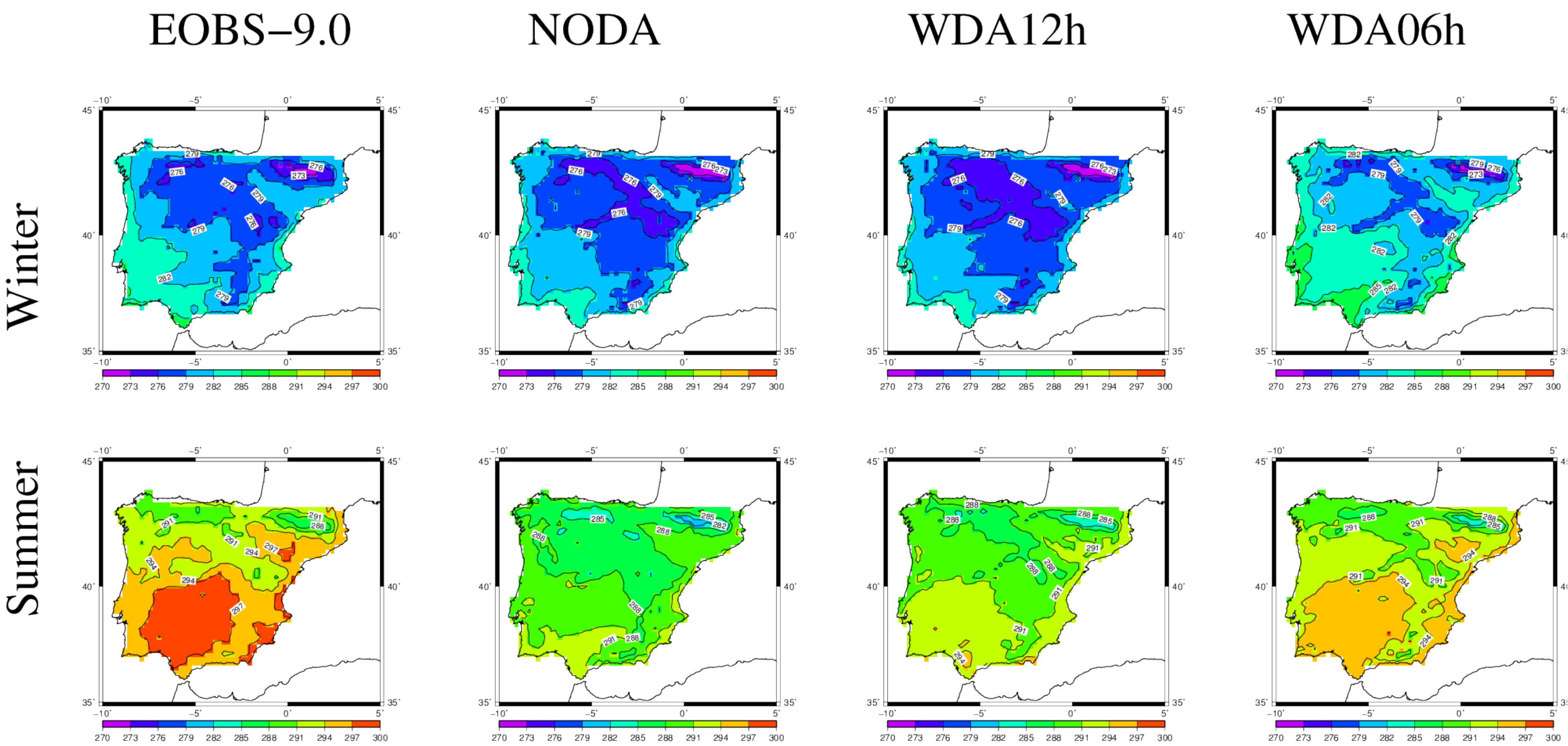


GTOPO30



IPRA

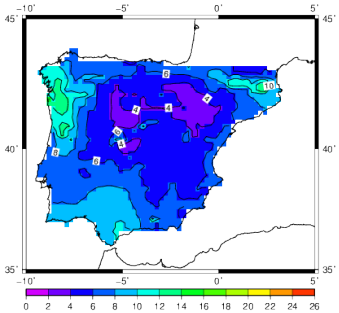
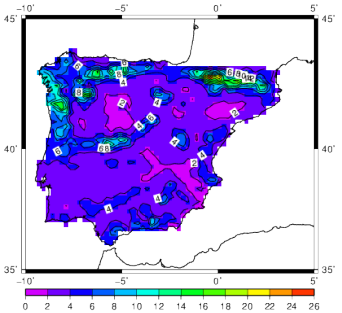
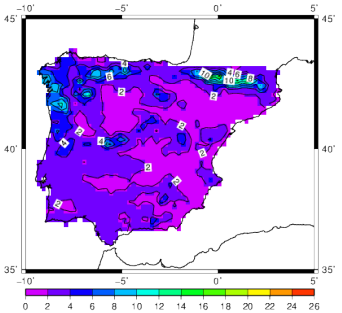
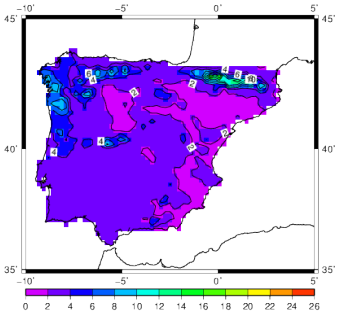
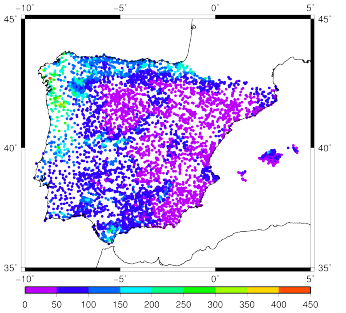
Seasonal averages - T2



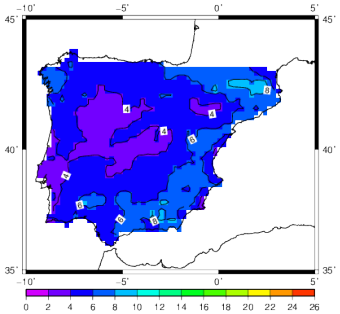
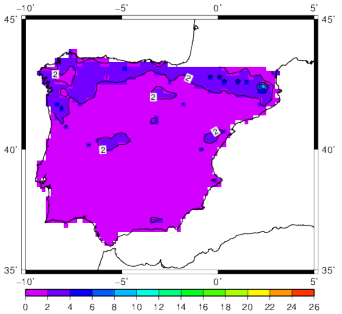
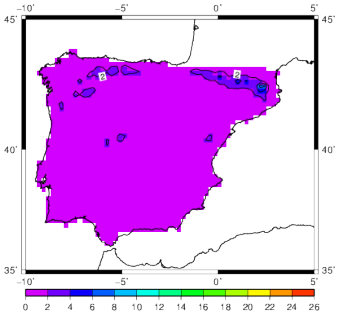
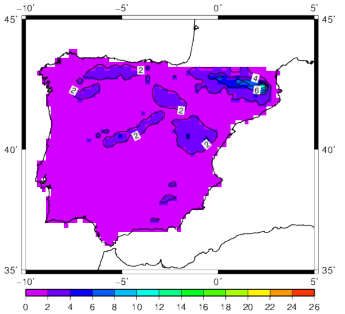
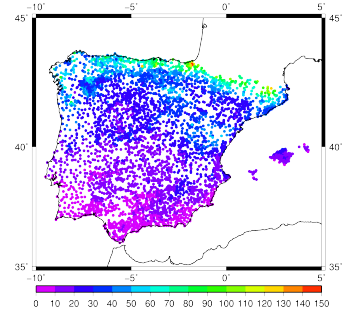
Seasonal averages - PRE

EOBS−9.0 | NODA | WDA12h | WDA06h | AEMET atlas

Winter

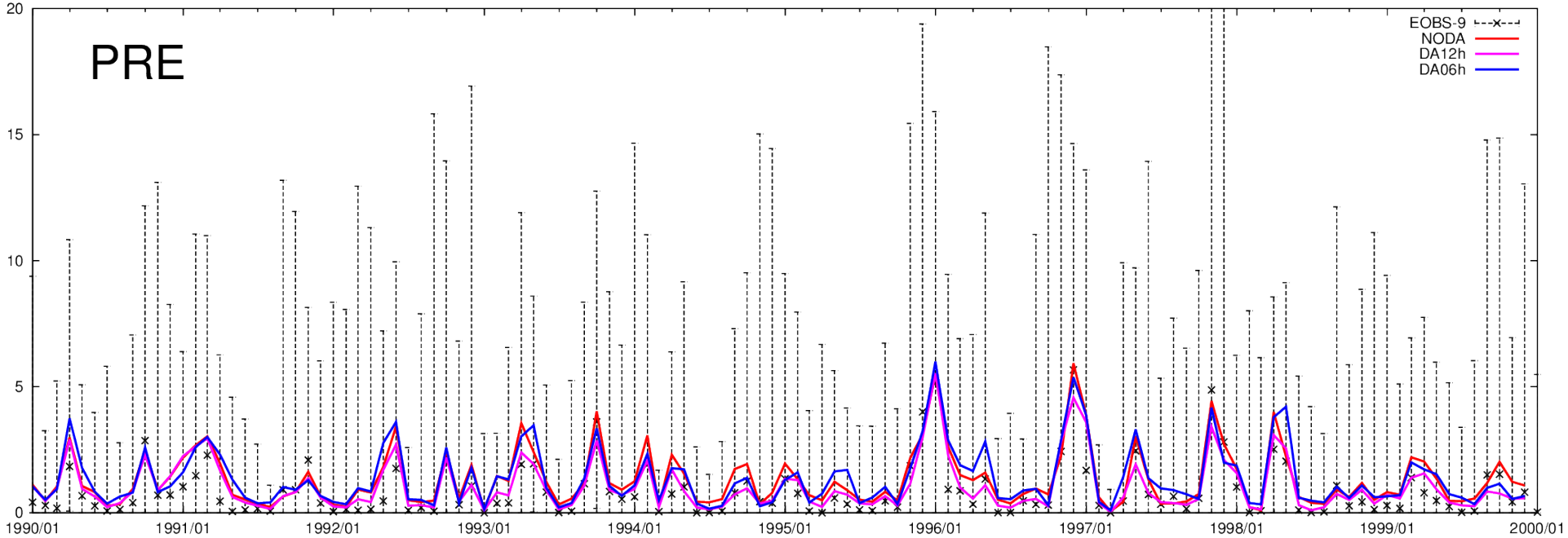
Summer

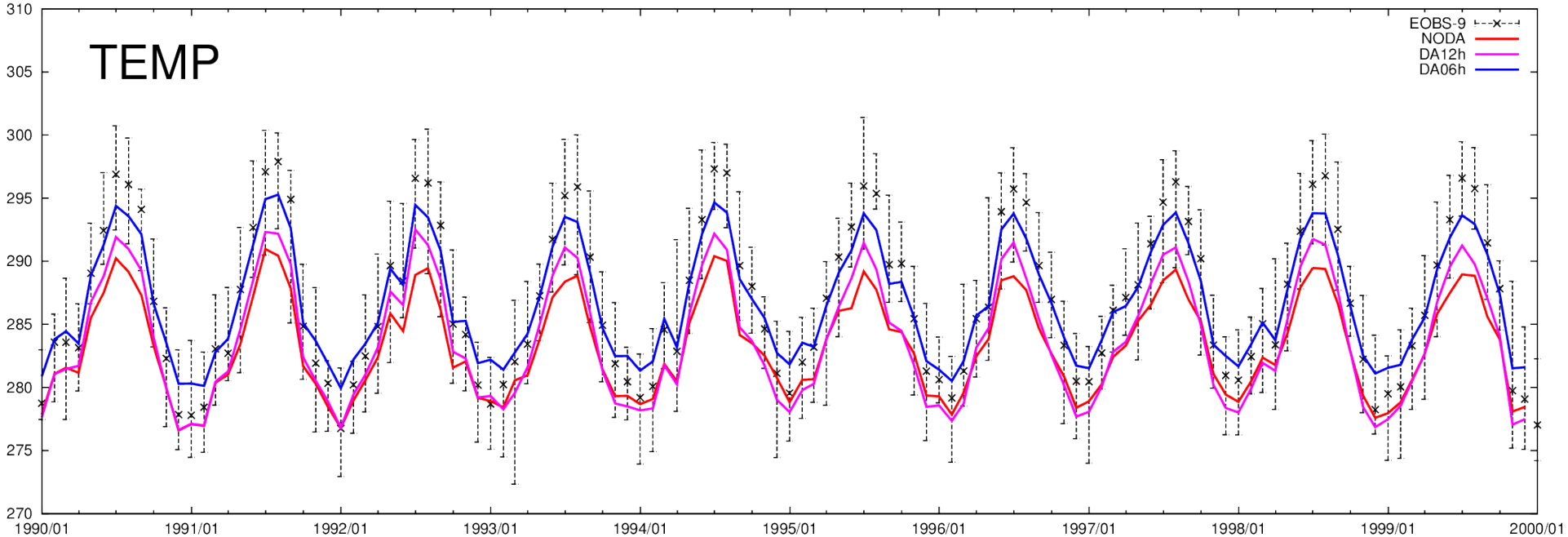






Daily!

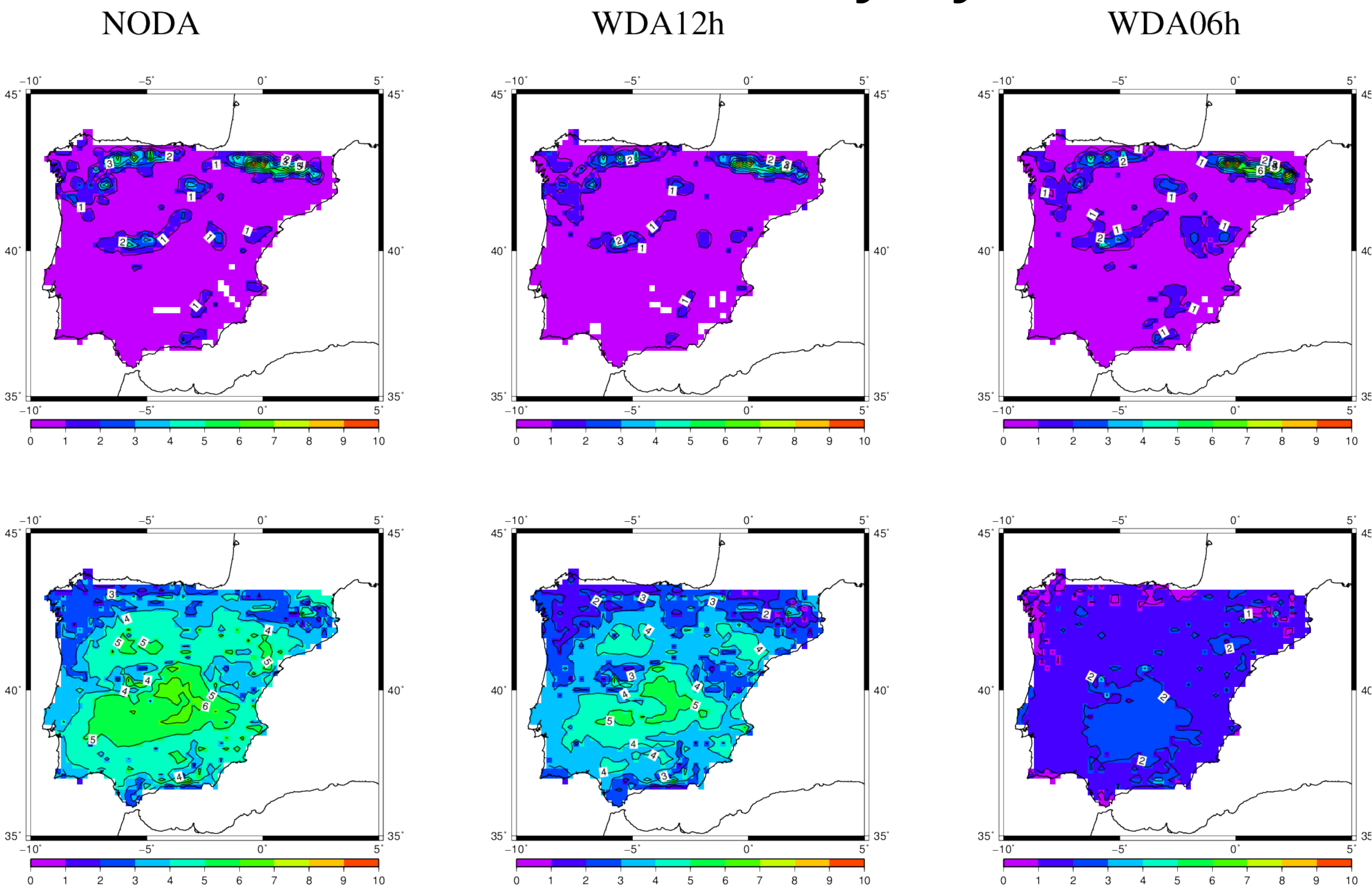
Monthly!

Area-averaged monthly medians





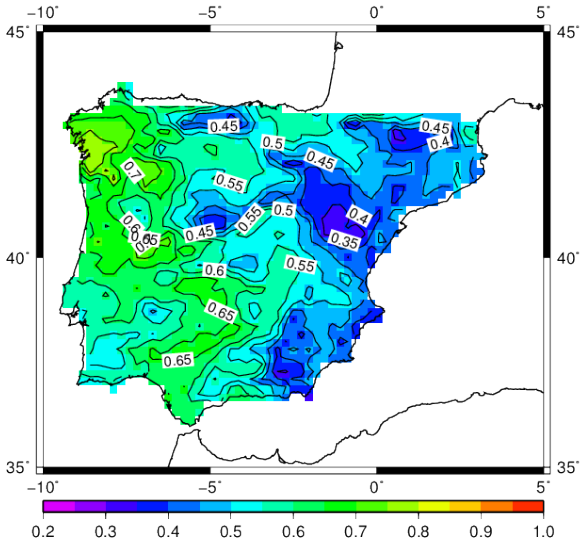
RMSE of the daily cycle

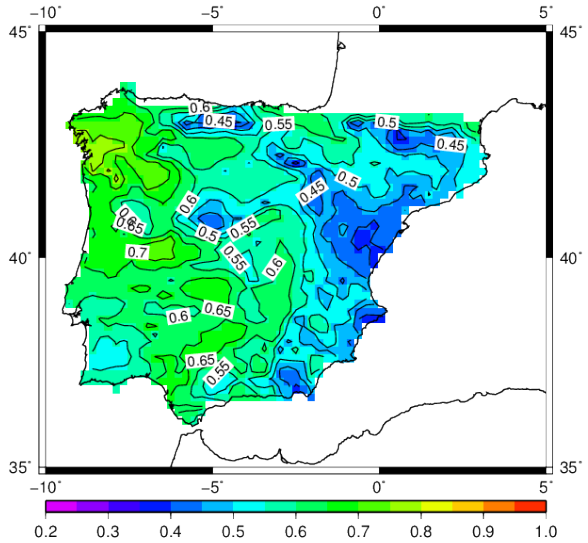


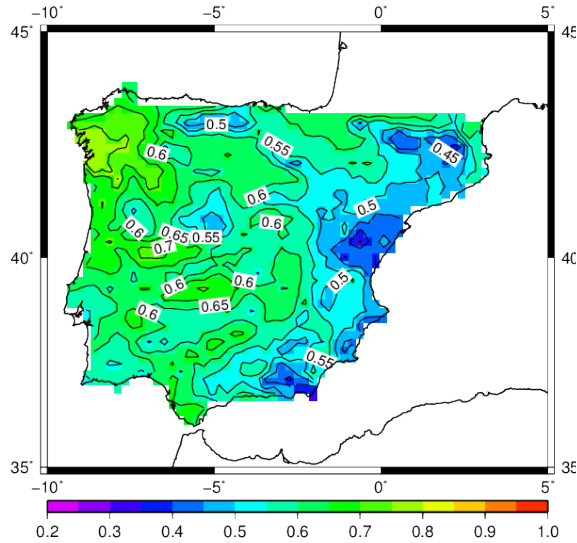
Correlation coefficient of anomalies

NO DA | DA–12h | DA–06h

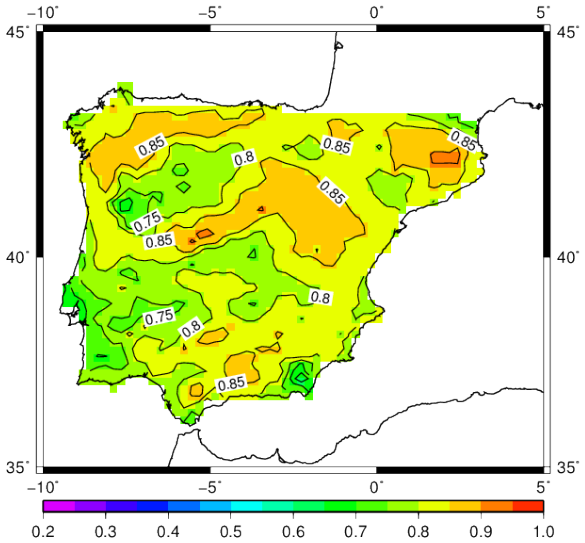
PRE

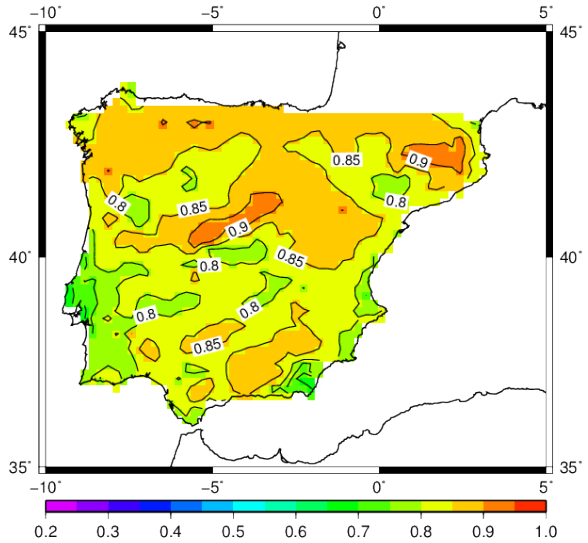


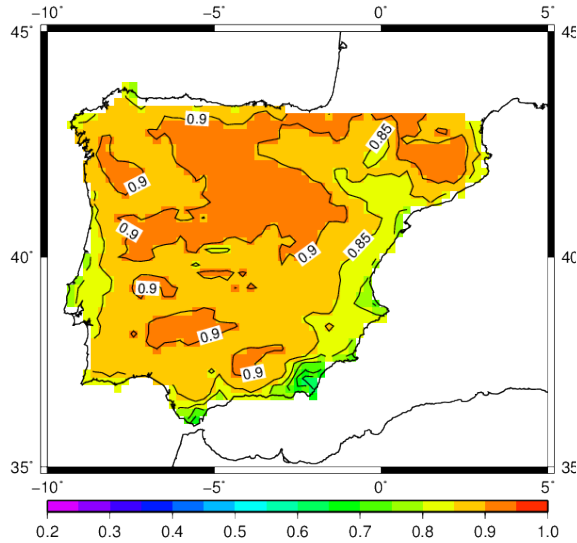




TEM



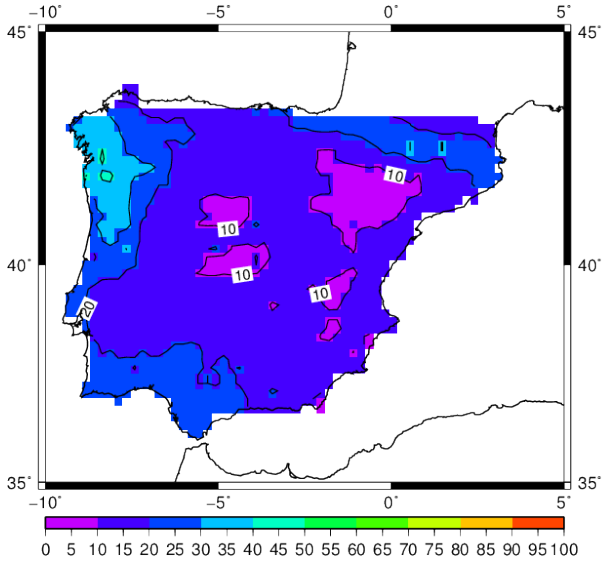




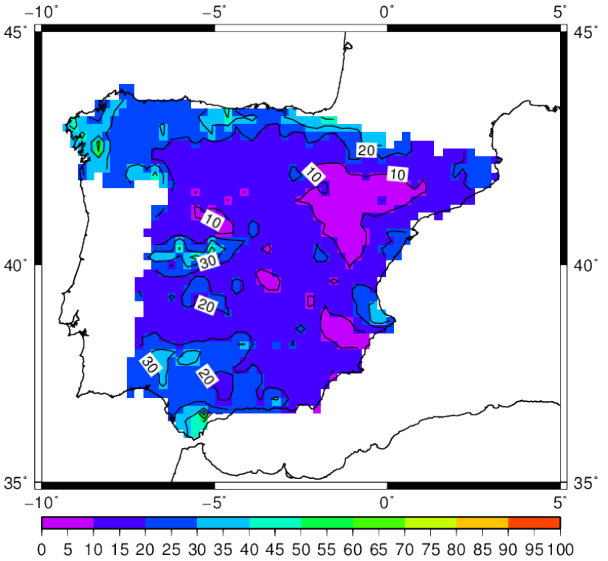
Anomalies after removing daily seasonal cycle defined by low-pass filtering daily seasonal cycle.
Significance computed by means of a Monte Carlo test that takes into account autocorrelation. Points reject H0 at 99.9%

P97.5 of daily precipitation - Winter

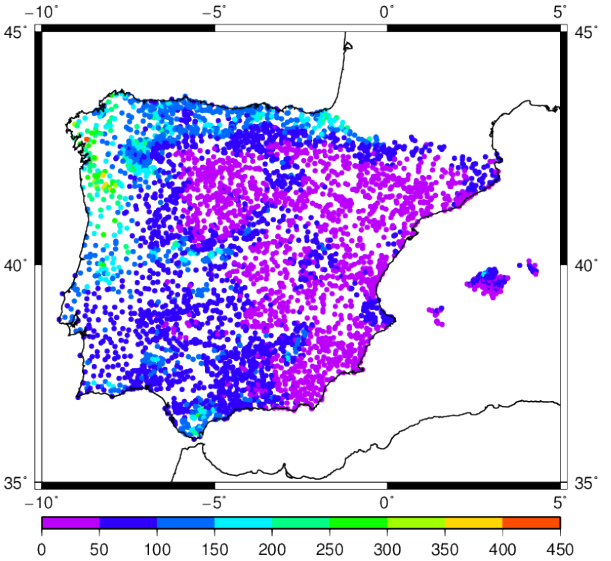
EOBS



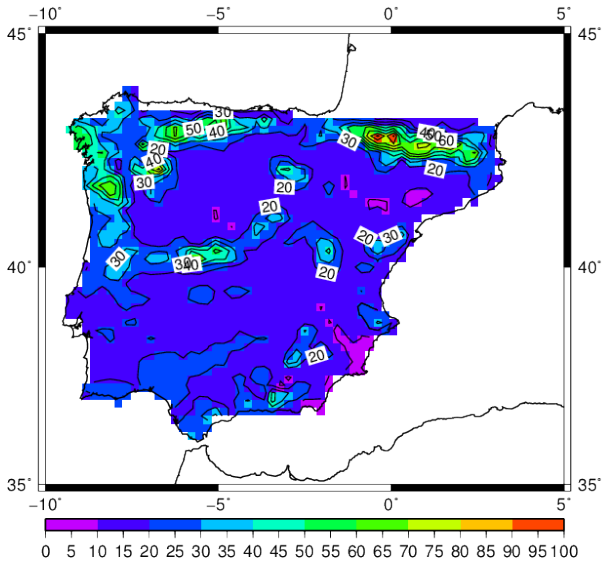
SP02



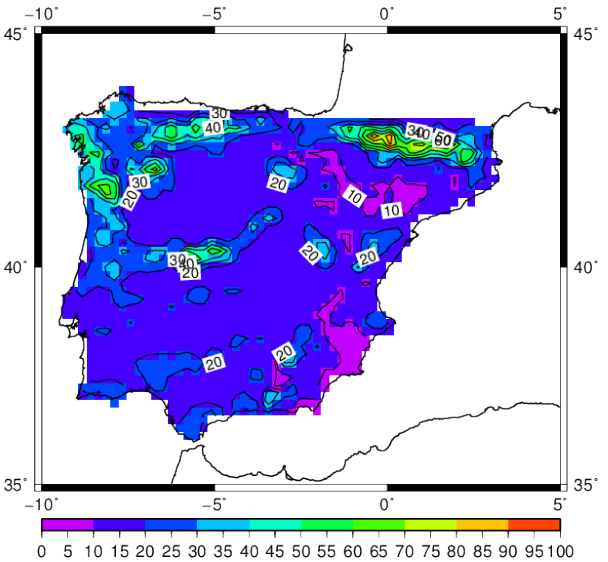
AEMET–ATLAS



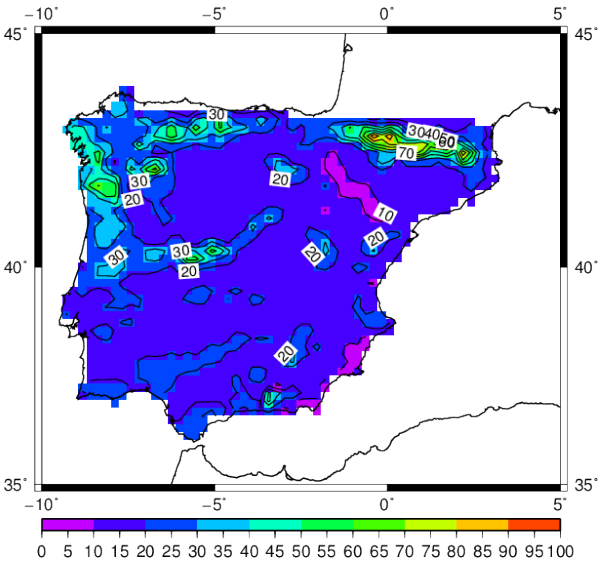
NODA



WDA12h

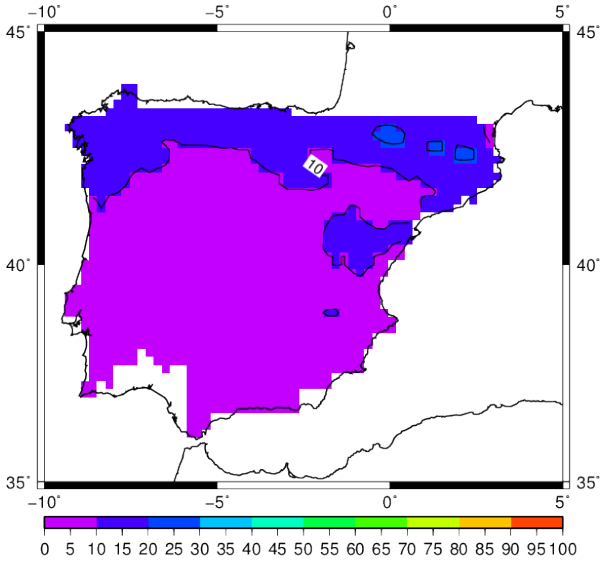


WDA06h

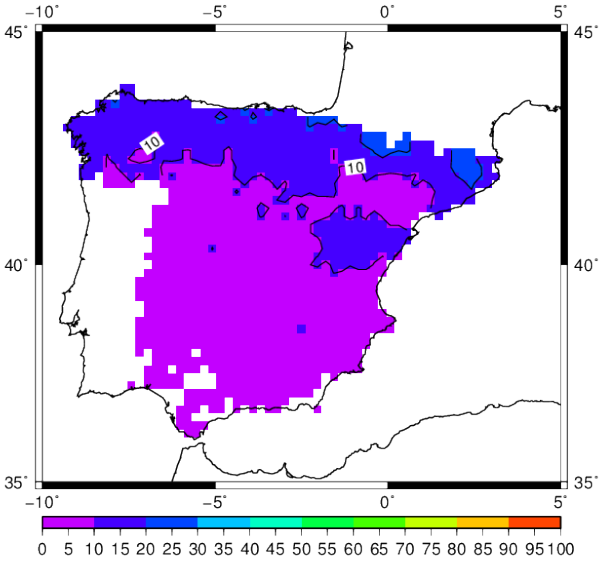


P97.5 of daily precipitation - Summer

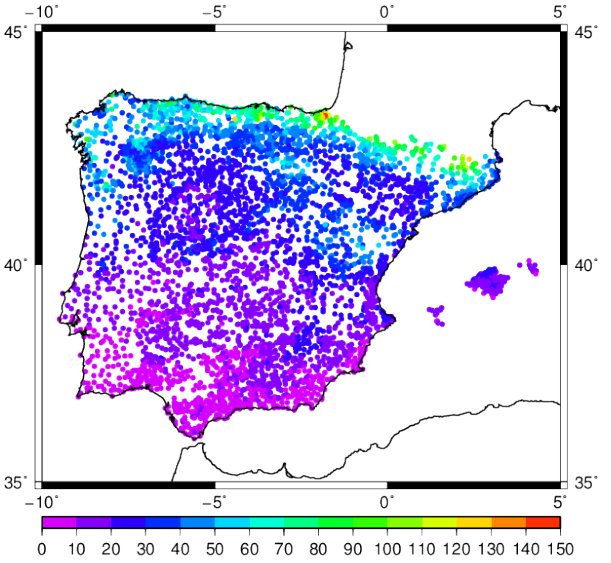
EOBS



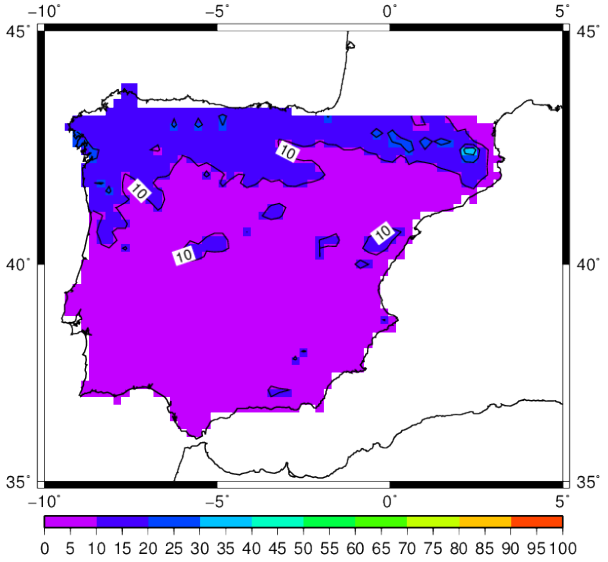
SP02



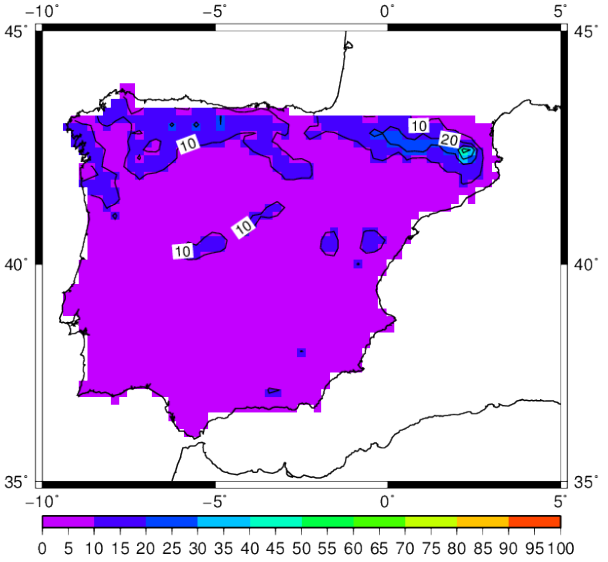
AEMET–ATLAS



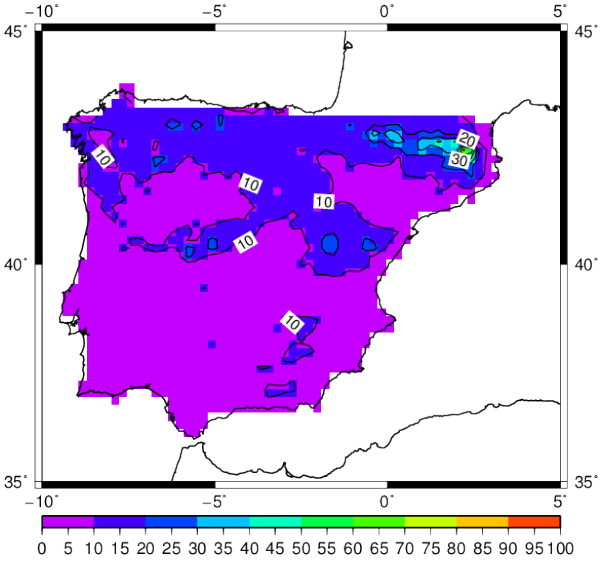
NODA



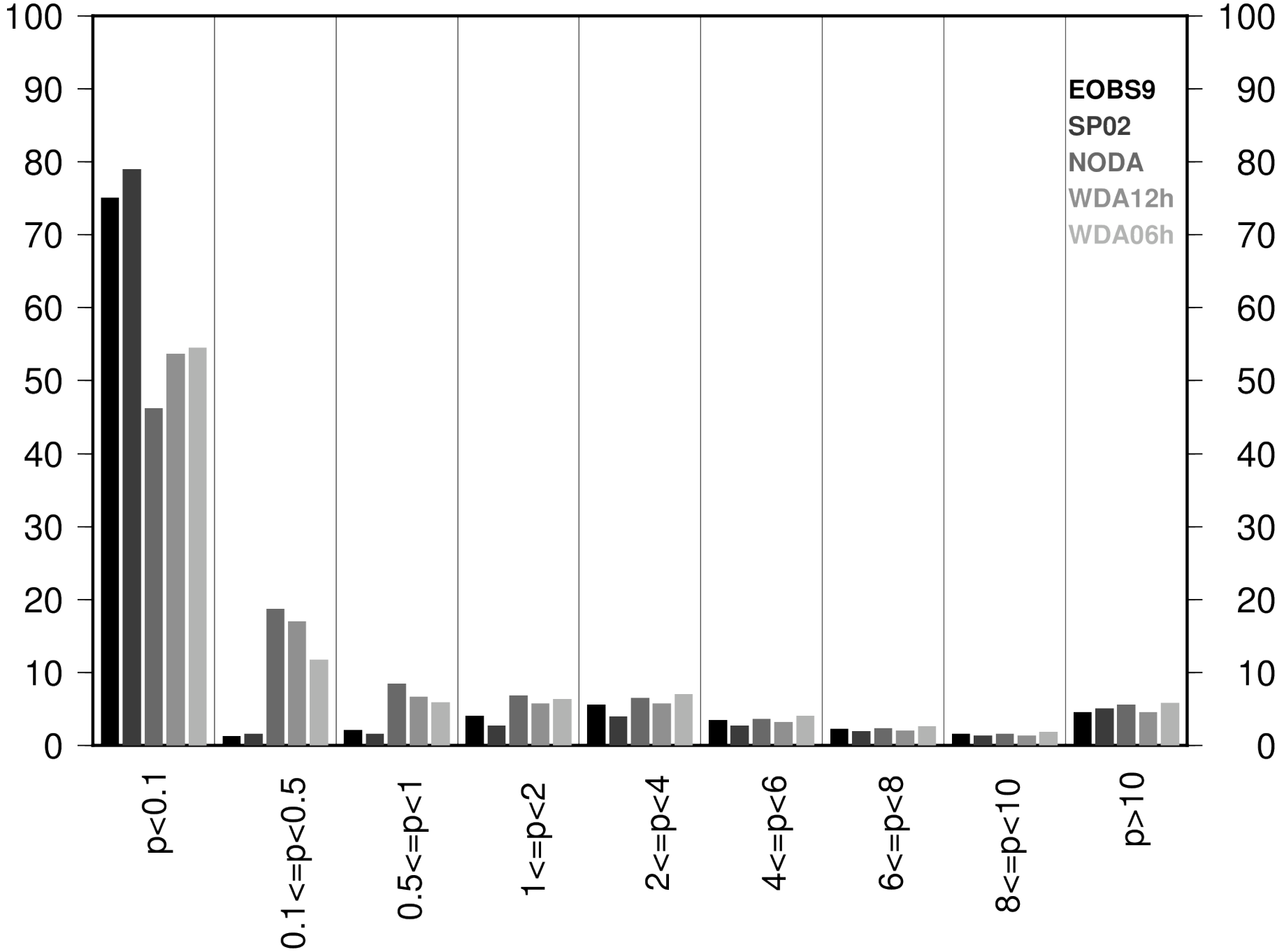
WDA12h



WDA06h



Histograms of daily precipitation



Value added over ERA40

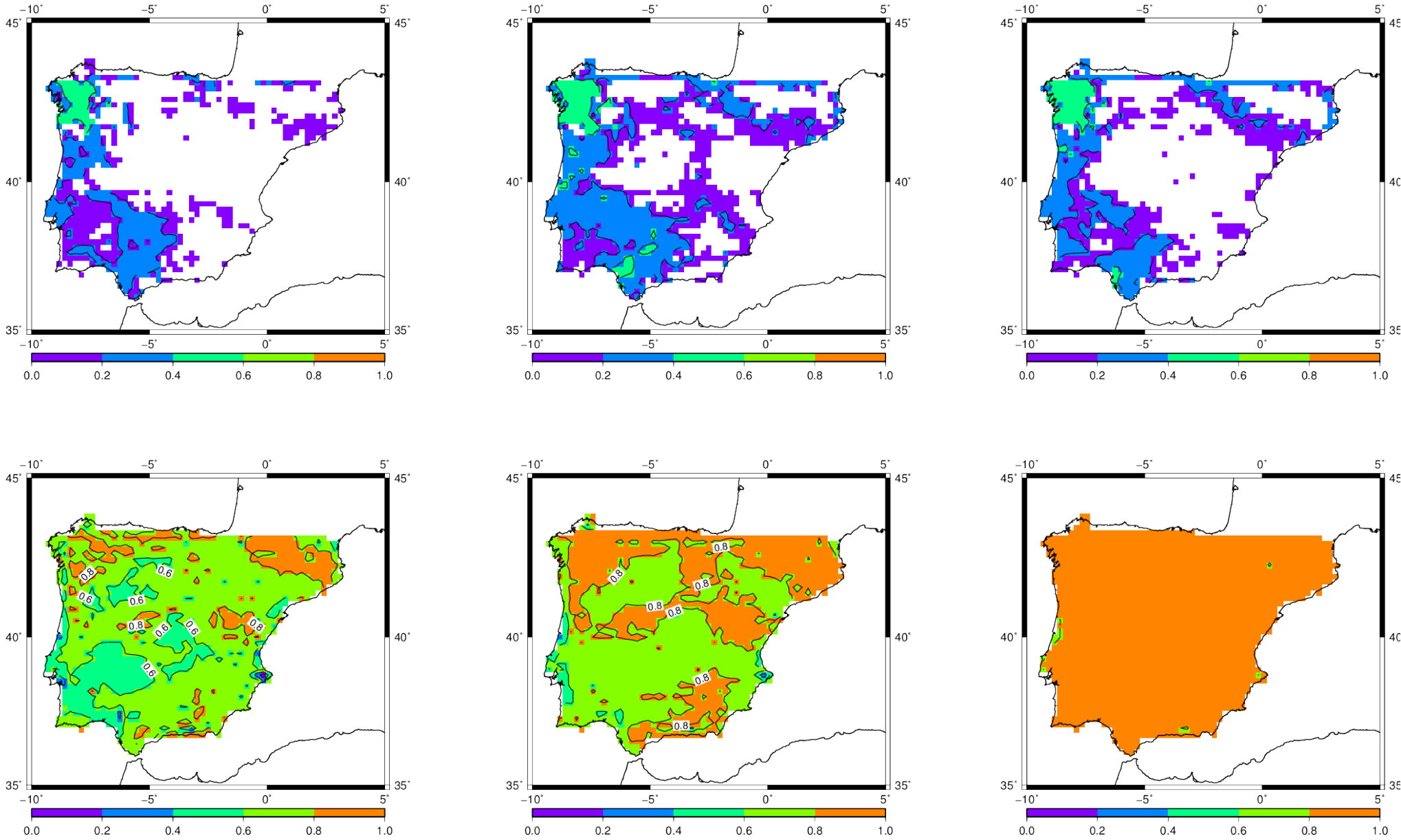
NO–DA

WDA–12h

WDA–06h

PRE

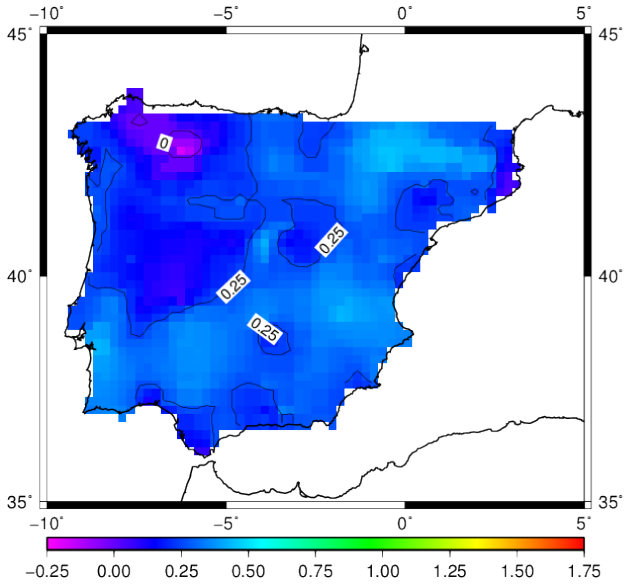
TEMP



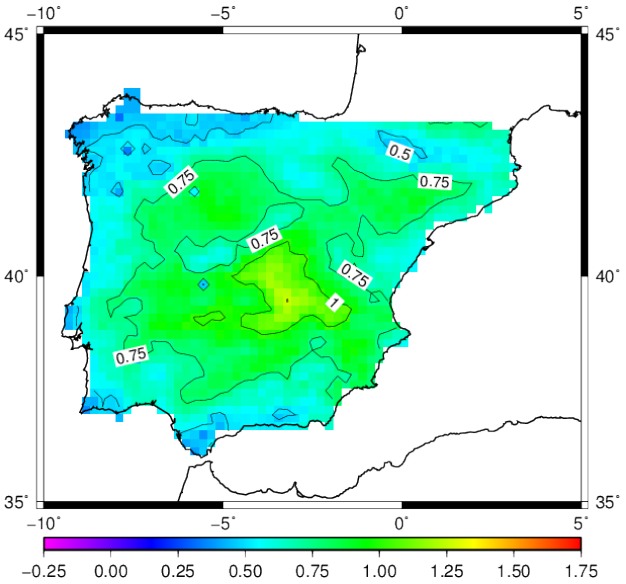
MSE-based skill score $VA = 1 - \frac{MSE_{IPvsEOBS}}{MSE_{ERAvsEOBS}}$

Long-term (1960-2002) T2m trends

EOBS

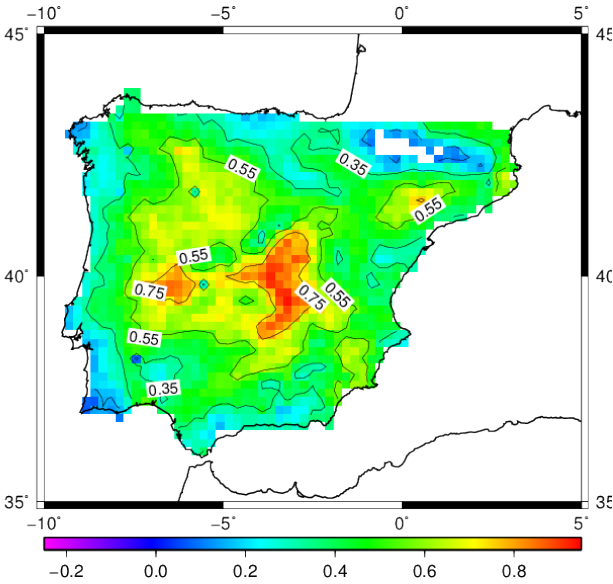


IPRA06



Too high

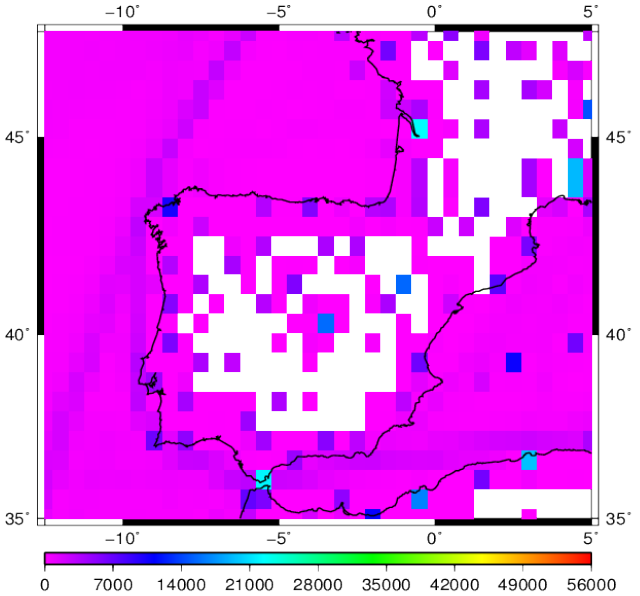
IPRA06–EOBS



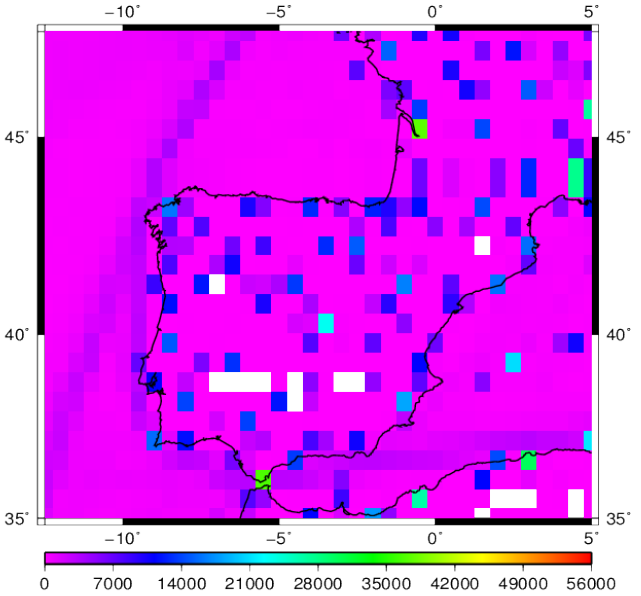
Signifance of difference: 95%

of observations available

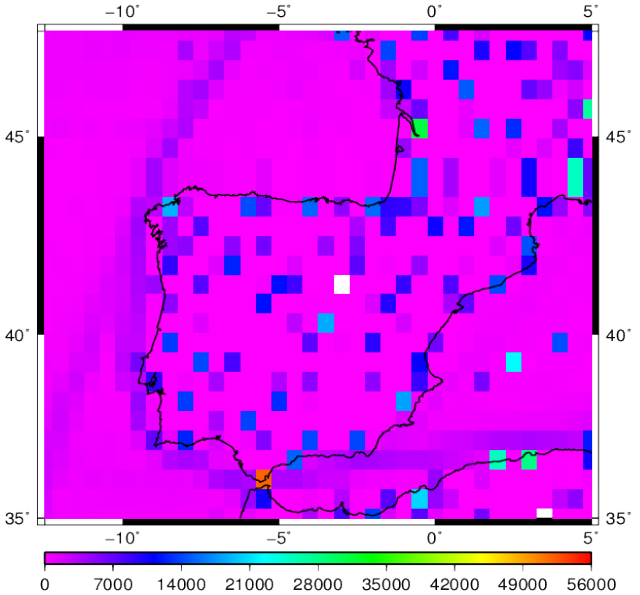
1960–1969



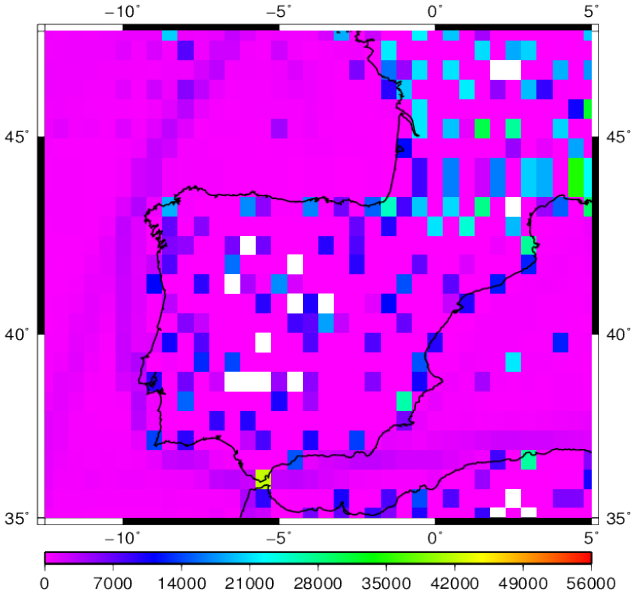
1970–1979



1980–1989

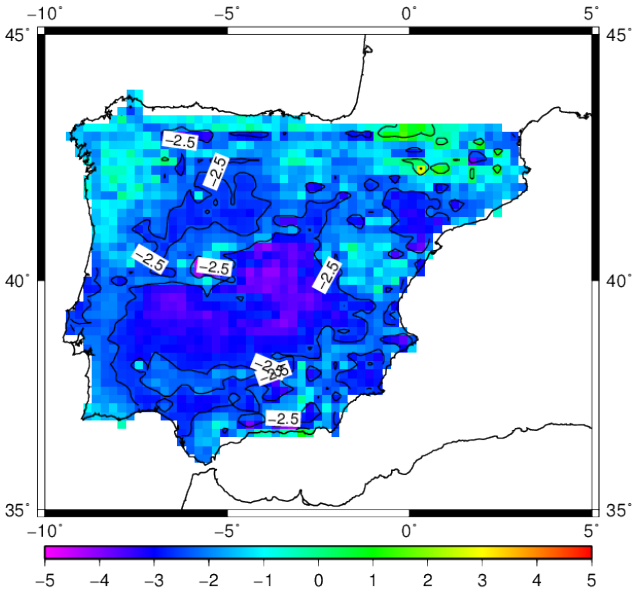


1990–1999

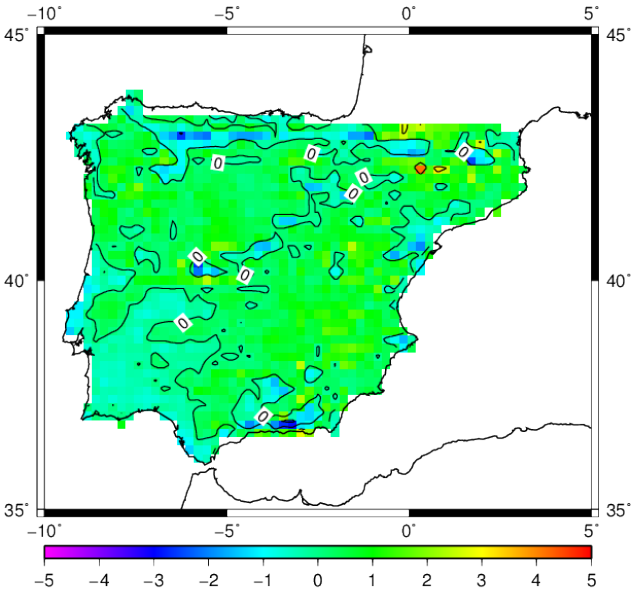


Time-varying T2m biases (WDA06h)

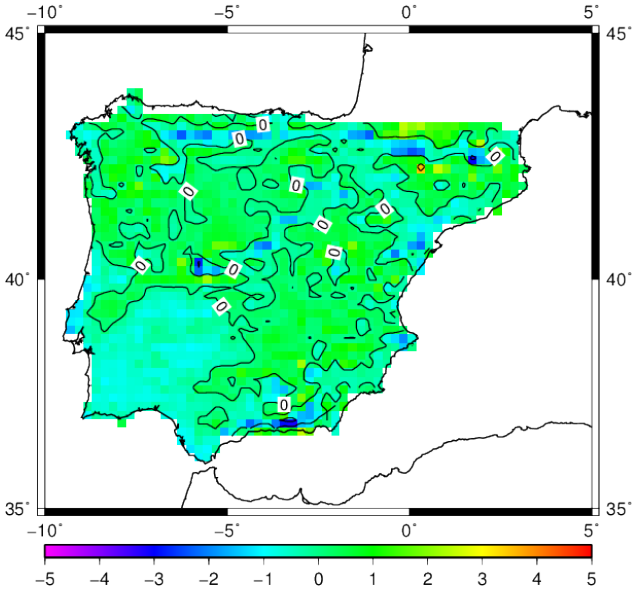
1960–1969



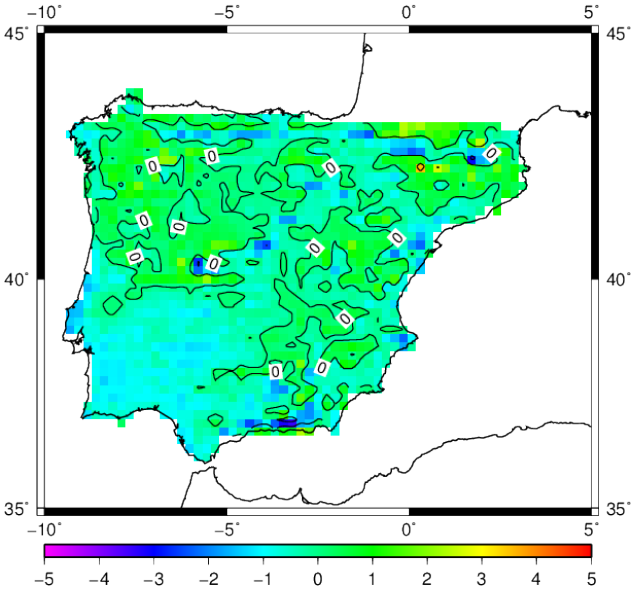
1970–1979



1980–1989



1990–1999



Computational resources

- ECMWF – Special Project ESIPRA
  - ecgate batch system for MARS requests
  - HPCF batch system for MARS requests and WRFDA
  - AIX, xlf, LoadLeveler, bash support *scarce*
- I2basque cluster
  - Linux, ifort, PBS
- RES project. MAGERIT.
  - AIX, xlf
- bash script for handling the integration: 24 sheets of commands and comments

Technical issues I

- WRF is quite portable
  - AIX-xlf (MAGERIT, ECMWF HPCF)
  - Linux-ifort (I2basque, arina-UPV/EHU, …)
- WRFDA is not that portable
  - AIX-xlf (~~MAGERIT~~, ~~ECMWF HPCF~~ serial only)
  - Linux-ifort (I2basque, arina-UPV/EHU, …)
- Parts of the code run sequentially
  - WPS-ungrib.exe, DA in ECMWF
  - LoadLeveler – ECMWF – nightmare!

Technical issues (II)

- IC/BC processing: ±2'10” every BC (6h). MARS access and retrieval
- OBS processing: ±2'00” extra MARS work (Δt≈100%!!) each analysis (6h/12h).
- WRF/WRFDA runs:
  - 13 months for the BE covariances
  - 10 years WRF run for NODA
  - 10 years WRF+DA run for WRDA12. ±23' per 24h run (3' DA and 20' WRF 24h)
  - 40+ years WRF+DA run for WRFDA06. ±13' per 12h run (3' DA and 10' WRF 12h)
- Total hours in I2Basque ONLY: 448 10^3 hours 2011-2013
- Output:
  - Archive: Over 11 Tb storing raw output without on-the-fly postprocessing (and ***losing***) data.
  - Still, you also need to postprocess the 11Tb outputs:
    - Vertical interpolation to P levels
    - Conversion of accumulated P to fluxes..

Conclusions

- We have performed an experiment over the Iberian Peninsula with a long integration (+ 3DVAR DA!!) using WRF and WRFDA
- For MOST of the skill scores analyzed, the integration using 3DVAR 6hourly works better than the one using 3DVAR 12h and better than the one using NO-DA.
- However, using 3DVAR is computationally more complex and expensive.
  - Getting access to observations implies writing proposals for ECMWF
  - Integrating the project requires access to RES.
  - Storing the results implies having access to a lot of hard disk.
- A publication is currently quite advanced and there are still several open questions:
  - Added value at different scales
  - What happens over sthe ocean?
  - Can we infer something from analysis increments?
- From the beginning our aim was to share the data openly through WWW, but just a subset will be open (very likely just sftp). Technicalities:
  - It is over 11Tb now.
  - Bandwidth at UPV/EHU? People managing the network will tell us...
  - THREDDS is not obvious for me/us. We are not system managers.
  - No money for technical suppport now.